Journal of Computational and Applied Mathematics 11 (1984) 83–94
North-Holland

Eigenfunctions and optimal orbits

R.A. CUNINGHAME-GREEN
Department of Mathematics, University of Birmingham, U.K.

R.E. BURKARD
Institut für Mathematik, Technische Universität Graz, A-8010 Graz, Austria

Received 3 January 1984

Abstract: For the algebraic structure ($\mathbb{R}$, max, +) we study the continuous analogue of the eigenvector-eigenvalue problem and relate it to a minimal-cost orbit problem. An explicit solution is given for the concave-quadratic case.

Keywords: Eigenfunctions, minimax algebra, least-cost orbit.

1. Introduction

For the algebraic structure ($\mathbb{R}$, $\oplus$, $\otimes$), sometimes called *max-algebra,* where, for $x, y, \in \mathbb{R}$:

$$x \oplus y = \max(x, y), \qquad x \otimes y = x + y, \tag{1.1}$$

we may pose analogues of the problems of classical linear algebra for a given matrix A:

The *eigenvector-eigenvalue problem:*

$$\text{solve} \quad A \otimes \boldsymbol{\xi} = \lambda \otimes \boldsymbol{\xi} \tag{1.2}$$

The *linear equations problem:*

$$\text{solve} \quad A \otimes \boldsymbol{\xi} = \boldsymbol{\eta}. \tag{1.3}$$

For a discussion of such structures and formulations, and their application to machine-scheduling, shortest-path, Boolean and other problems, see [1,2] and references cited therein.

In particular, (1.3) leads to a discussion of the dual algebra ($\mathbb{R}$, $\oplus'$, $\otimes'$), sometimes called *min-algebra* where for $x, y, \in \mathbb{R}$:

$$x \oplus' y = \min(x, y), \qquad x \otimes' y = x + y. \tag{1.4}$$

If A^* denotes the negative transposed of A and primes denote the use of min-algebra, then (1.3) if soluble always possesses a *principal solution*

$$\boldsymbol{\xi} = A^* \otimes' \boldsymbol{\eta} \tag{1.5}$$

and the triple product

$$A \otimes (A^* \otimes' \eta) \tag{1.6}$$

is in any case a Chebychev-best under-approximation of η from the column-space of A.

2. The continuous analogue

The results discussed in Section 1 all have continuous analogues, in which the matrix A is replaced by a function

$$A(x, y) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \tag{2.1}$$

whilst the vectors ξ, η are replaced by functions

$$f(x), g(x) : \mathbb{R}^n \to \mathbb{R}. \tag{2.2}$$

The relationship between (1.3) and (1.5) then becomes a generalisation of Fenchel's theory of conjugate functions, whilst (1.6) becomes a *generalised convexification* of a given function. Related ideas have been developed by e.g. Brøndsted [3] and Moreau [4].

In the present paper, we shall study the continuous analogue of (1.2). Now (1.2) itself asks, given an $(n \times n)$ matrix $A = [a_{ij}]$ ($a_{ij} \in \mathbb{R}$, $i, j = 1, \ldots, n$), to find $\xi = [\xi_j] \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ such that

$$\max_{j=1,\ldots,n} (a_{ij} + \xi_j) = \lambda + \xi_i, \quad i = 1, \ldots, n. \tag{2.3}$$

The continuous analogue of this is: given $A(x, y)$ of (2.1), find an *eigenfunction*

$$f : \mathbb{R}^n \to \mathbb{R} \tag{2.4}$$

and eigenvalue $\lambda \in \mathbb{R}$ such that

$$\max_{y \in \mathbb{R}^n} (A(x, y) + f(y)) = \lambda + f(x) \quad \forall x \in \mathbb{R}^n. \tag{2.5}$$

Just as (1.2) is related to the classical shortest-path problem [1], we shall show that this continuous analogue is related to the following *optimal orbit problem:*

Points in $\mathbb{R}^n$ represent the states of some system and a function

$$A(x, y) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{2.6}$$

is given, representing the profit (if negative, the loss) associated with a transition from state x to state y. If the profit of a sequence of transitions is the sum of their individual profits, find a strategy for moving from given state η to given state ξ in a (possibly infinite) sequence of transitions, at maximum total profit.

3. Assumptions, norms and notations

All derivatives used will be Fréchet derivatives. We assume that:

(1) A has derivatives of all necessary orders.

(2) For each $x \in \mathbb{R}^n$, the maximisation with respect to y in (2.5) is achieved at a unique point (depending on x), say at $\theta(x)$.

(3) $\boldsymbol{\theta}$ has derivatives of all necessary orders.
(4) $\boldsymbol{\theta}$ is bounded.
Hence there is a compact set $\mathscr{C} \subset \mathbb{R}^n$ (which may be taken arbitrarily large) whose interior contains the range of $\boldsymbol{\theta}$.

In the ensuing arguments, we shall employ a number of norms.

For $y \in \mathbb{R}^n$, $\|y\|$ denotes the Euclidean norm and then for functions such as $\boldsymbol{\theta}: \mathbb{R}^n \to \mathbb{R}^n$, $\|\boldsymbol{\theta}\|$ denotes $\mathrm{Sup}_{z \in \mathscr{C}} \|\boldsymbol{\theta}(z)\|$. For a constant matrix A, $\|A\|$ denotes $\sup_{\|x\|=1} \|Ax\|$; then for a matrix $M = M(z_1, \ldots, z_r)$ whose components are functions of $z_1, \ldots, z_r \in \mathbb{R}^n$ we define

$$\|M\| = \sup_{z_1, \ldots, z_r \in \mathscr{C}} \|M(z_1, \ldots, z_r)\|.$$

In particular, since we shall not distinguish between the Fréchet derivative B of a function $\boldsymbol{\beta}: \mathbb{R}^m \to \mathbb{R}^n$ and the matrix by which B is presented, we have thereby assigned a norm to such derivatives. When $\boldsymbol{\beta}$ maps to $\mathbb{R}$, B is of course the row-vector $(\nabla \boldsymbol{\beta})$.

For a function of $2n$ variables we shall use ∇_1 and ∇_2 to denote the taking of gradients with respect to the first and the second argument positions, respectively. To avoid ambiguity we write e.g.

$$(\nabla f)(\boldsymbol{\theta}(x)) \quad \text{or} \quad (\nabla_1 A)(y, \boldsymbol{\theta}(x))$$

to mean: first the gradients are formed; then the arguments are set. The $\boldsymbol{\theta}$, x and y take no part in the differentiation.

For clarity we shall occasionally use a dot $\cdot$ to indicate that a matrix multiplication is intended, unless this is obvious. All zeros (matrix, vector or scalar) will be indifferently denoted 0.

4. Formal solution

Now (2.5) and Assumption (2) imply on the one hand that

$$A(x, \boldsymbol{\theta}(x)) + f(\boldsymbol{\theta}(x)) = \lambda + f(x) \quad \forall\, x \in \mathscr{C}. \tag{4.1}$$

On the other hand $y = \boldsymbol{\theta}(x)$ is a maximising, therefore stationary, point of $A(x, y) + f(y)$ quâ function of y, so

$$(\nabla_2 A)(x, \boldsymbol{\theta}(x)) + (\nabla f)(\boldsymbol{\theta}(x)) = 0 \quad \forall\, x \in \mathscr{C}. \tag{4.2}$$

Taking the gradient of (4.1) quâ function of x:

$$(\nabla_1 A)(x, \boldsymbol{\theta}(x)) + [(\nabla_2 A)(x, \boldsymbol{\theta}(x))]\Theta(x) + [(\nabla f)(\boldsymbol{\theta}(x))]\Theta(x) = (\nabla f)(x), \tag{4.3}$$

where $\Theta(x)$ is the Fréchet derivative of $\boldsymbol{\theta}$ at x. But right-multiplying (4.2) by $\Theta(x)$ simplifies (4.3) to

$$(\nabla_1 A)(x, \boldsymbol{\theta}(x)) = (\nabla f)(x) \quad \forall\, x \in \mathscr{C}. \tag{4.4}$$

In particular, taking $\boldsymbol{\theta}(x) \in \mathscr{C}$ in the role of x, (4.4) implies

$$(\nabla_1 A)(\boldsymbol{\theta}(x), \boldsymbol{\theta}^2(x)) = (\nabla f)(\boldsymbol{\theta}(x)) \quad \forall\, x \in \mathscr{C}, \tag{4.5}$$

where $\boldsymbol{\theta}^2(x) = \boldsymbol{\theta}(\boldsymbol{\theta}(x))$.

Comparing (4.2) and (4.5) yields as a necessary condition on $\boldsymbol{\theta}$:

$$(\nabla_1 A)(\boldsymbol{\theta}(\boldsymbol{x}), \boldsymbol{\theta}^2(\boldsymbol{x})) + (\nabla_2 A)(\boldsymbol{x}, \boldsymbol{\theta}(\boldsymbol{x})) = 0 \quad \forall\, \boldsymbol{x} \in \mathscr{C}. \tag{4.6}$$

Expression (4.6) is an implicit relationship which is (in principle) soluble for $\boldsymbol{\theta}$, given A, and then we may integrate (4.4) to obtain an eigenfunction $f(\boldsymbol{x})$. (Notice that (2.5) only requires f to be known to within an additive constant of integration.) Then for any arbitrary $\boldsymbol{x} \in \mathscr{C}$, (4.1) gives

$$\lambda = A(\boldsymbol{x}, \boldsymbol{\theta}(\boldsymbol{x})) + f(\boldsymbol{\theta}(\boldsymbol{x})) - f(\boldsymbol{x}). \tag{4.7}$$

Theorem 4.1. *If $\boldsymbol{\theta}$ has a fixed point $\boldsymbol{\xi} \in \mathscr{C}$, then*

$$\lambda = A(\boldsymbol{\xi}, \boldsymbol{\xi}) = \max_{z \in \mathscr{C}} A(z, z).$$

Proof. From (2.5),

$$\lambda + f(z) \geqslant A(z, z) + f(z) \quad \forall\, z \in \mathscr{C}.$$

Hence $\lambda \geqslant A(z, z)$, but from (4.7) $\lambda = A(\boldsymbol{\xi}, \boldsymbol{\xi})$. □

5. Solubility of (4.6)

We now produce a sufficient condition for the existence of solutions $\boldsymbol{\theta}$ to (4.6), having a fixed point. First, some notation.

Let

$$\Gamma = \Gamma(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix}$$

be a $(2n \times 2n)$ functional matrix partitioned into four $(n \times n)$ matrices. The (ij)th element of Γ is

$$(\partial_i \partial_j A)(\boldsymbol{u}, \boldsymbol{v}) \quad \text{for } i = 1, \ldots, n;\ j = 1, \ldots, 2n,$$

$$(\partial_i \partial_j A)(\boldsymbol{w}, \boldsymbol{u}) \quad \text{for } i = n+1, \ldots, 2n;\ j = 1, \ldots, 2n,$$

where ∂_i denotes the partial derivative of a function with respect to its ith argument position.

Theorem 5.1. *If $(\Gamma_{11} + \Gamma_{22})$ is nonsingular for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathscr{C}$, and also for some constant $\rho (0 < \rho < \frac{1}{2})$ there holds*

$$\max\left(\|(\Gamma_{11} + \Gamma_{22})^{-1} \Gamma_{12}\|, \|(\Gamma_{11} + \Gamma_{22})^{-1} \Gamma_{21}\| \right) < \rho, \tag{5.1}$$

then unique $\boldsymbol{\theta} : \mathscr{C} \to \mathscr{C}$ exists, satisfying (4.6) and such that its derivative satisfies $\|\Theta\| < 1$.

Proof. Let $\phi : \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ be the function implicitly defined by saying that

$$\boldsymbol{u} = \boldsymbol{\phi}(\boldsymbol{v}, \boldsymbol{w}) \iff (\nabla_1 A)(\boldsymbol{u}, \boldsymbol{v}) + (\nabla_2 A)(\boldsymbol{w}, \boldsymbol{u}) = 0. \tag{5.2}$$

The nonsingularity of $(\Gamma_{11} + \Gamma_{22})$ guarantees by the implicit function theorem that ϕ is well-defined, and its derivative Φ exists and is given by

$$\Phi(\boldsymbol{v}, \boldsymbol{w}) = -(\Gamma_{11} + \Gamma_{22})^{-1} [\Gamma_{12}, \Gamma_{21}] \quad \text{with } \boldsymbol{u} = \boldsymbol{\phi}(\boldsymbol{v}, \boldsymbol{w}) \text{ in } \Gamma. \tag{5.3}$$

Now choose σ such that $2\rho < \sigma < 1$. Take any function $\boldsymbol{\theta}_0 : \mathscr{C} \to \mathscr{C}$ whose derivative Θ_0 satisfies $\|\Theta_0\| < \sigma$. Construct a sequence of functions $\boldsymbol{\theta}_r : \mathscr{C} \to \mathscr{C}$ $(r = 0, 1, \ldots)$ by defining

$$\boldsymbol{\theta}_{r+1}(\boldsymbol{x}) = \boldsymbol{\phi}(\boldsymbol{\theta}_r^2(\boldsymbol{x}), \boldsymbol{x}), \tag{5.4}$$

where $\boldsymbol{\theta}_r^2(\boldsymbol{x})$ denotes $\boldsymbol{\theta}_r(\boldsymbol{\theta}_r(\boldsymbol{x}))$ and $\boldsymbol{\phi}$ is defined by (5.2).

$$\text{Assume that } \|\Theta_r\| < \sigma \text{ for particular } r \geqslant 0. \tag{5.5}$$

From (5.4):

$$\Theta_{r+1}(\boldsymbol{x}) = \Phi(\boldsymbol{\theta}_r^2(\boldsymbol{x}), \boldsymbol{x})\begin{bmatrix} \Theta_r(\boldsymbol{\theta}_r(\boldsymbol{x}))\Theta_r(\boldsymbol{x}) \\ I \end{bmatrix}. \tag{5.6}$$

So by (5.3), (5.6):

$$\begin{aligned} \|\Theta_{r+1}\| &\leqslant \|(\Gamma_{11} + \Gamma_{22})^{-1}\Gamma_{12}\|\|\Theta_r\|^2 + \|(\Gamma_{11} + \Gamma_{22})^{-1}\Gamma_{21}\| \\ &< \rho(\sigma^2 + 1) \quad \text{(by (5.1), (5.5))} \\ &< \rho(\sigma + 1) \quad \text{(because } 0 < \sigma < 1) \\ &< \sigma \quad \text{(by choice of } \sigma). \end{aligned}$$

Hence by induction:

$$\|\Theta_r\| < \sigma, \tag{5.7}$$

where $0 < \sigma < 1$, for $r = 0, 1, \ldots$. Now consider another sequence of functions $\boldsymbol{\psi}_r : \mathscr{C} \to \mathscr{C}$ (with derivatives Ψ_r) constructed in the same way as the functions $\boldsymbol{\theta}_r$ but with a different starting function $\boldsymbol{\psi}_0$ satisfying $\|\Psi_0\| < \sigma$. Then by the Mean Value Theorem and (5.3), (5.4):

$$\begin{aligned} \|\boldsymbol{\theta}_{r+1} - \boldsymbol{\psi}_{r+1}\| &= \operatorname*{Sup}_{x \in \mathscr{C}} \|\boldsymbol{\phi}(\boldsymbol{\theta}_r^2(\boldsymbol{x}), \boldsymbol{x}) - \boldsymbol{\phi}(\boldsymbol{\psi}_r^2(\boldsymbol{x}), \boldsymbol{x})\| \\ &= \operatorname*{Sup}_{x \in \mathscr{C}} \|\Pi \cdot (\boldsymbol{\theta}_r^2(\boldsymbol{x}) - \boldsymbol{\psi}_r^2(\boldsymbol{x}))\|, \end{aligned} \tag{5.8}$$

where $\Pi = -(\Gamma_{11} + \Gamma_{22})^{-1}\Gamma_{12}$ with suitable arguments depending on $\boldsymbol{x}$, $\boldsymbol{\theta}_r$, $\boldsymbol{\psi}_r$. Also by the Mean Value Theorem:

$$\begin{aligned} \|\boldsymbol{\theta}_r^2(\boldsymbol{x}) - \boldsymbol{\psi}_r^2(\boldsymbol{x})\| &= \|\boldsymbol{\theta}_r(\boldsymbol{\theta}_r(\boldsymbol{x})) - \boldsymbol{\theta}_r(\boldsymbol{\psi}_r(\boldsymbol{x})) + \boldsymbol{\theta}_r(\boldsymbol{\psi}_r(\boldsymbol{x})) - \boldsymbol{\psi}_r(\boldsymbol{\psi}_r(\boldsymbol{x}))\| \\ &= \|\Theta_r(\boldsymbol{\eta})(\boldsymbol{\theta}_r(\boldsymbol{x}) - \boldsymbol{\psi}_r(\boldsymbol{x})) + \boldsymbol{\theta}_r(\boldsymbol{\psi}_r(\boldsymbol{x})) - \boldsymbol{\psi}_r(\boldsymbol{\psi}_r(\boldsymbol{x}))\| \end{aligned} \tag{5.9}$$

for suitable $\boldsymbol{\eta}$ depending on $\boldsymbol{x}$, $\boldsymbol{\theta}_r$ and $\boldsymbol{\psi}_r$.

Hence from (5.1), (5.7), (5.8), (5.9):

$$\|\boldsymbol{\theta}_{r+1} - \boldsymbol{\psi}_{r+1}\| < \rho(\sigma + 1)\|\boldsymbol{\theta}_r - \boldsymbol{\psi}_r\|$$

But

$$\rho(\sigma + 1) < \sigma \quad \text{(by choice of } \sigma) < 1. \tag{5.10}$$

Hence the transformation $\boldsymbol{\theta}_r \mapsto \boldsymbol{\theta}_{r+1}$ defined by (5.4) is a contraction mapping and therefore there exists a function $\boldsymbol{\theta}$ invariant under this transformation, i.e. satisfying

$$\boldsymbol{\theta}(\boldsymbol{x}) = \boldsymbol{\phi}(\boldsymbol{\theta}^2(\boldsymbol{x}), \boldsymbol{x}),$$

and so, by (5.2), satisfying (4.6). Moreover, the sequence of functions $\boldsymbol{\theta}_r$, converging pointwise to $\boldsymbol{\theta}$, satisfy $\|\Theta_r\| < \sigma$, so $\boldsymbol{\theta}$ satisfies $\|\Theta\| < 1$. Conversely, if $\boldsymbol{\psi}$ were any solution to (4.6) satisfying $\|\Psi\| < 1$, then taking $\sigma > \|\Psi\|$ and $\boldsymbol{\psi}_0 = \boldsymbol{\psi}$ in the above argument, we easily see that $\boldsymbol{\psi} = \boldsymbol{\theta}$. $\square$

Corollary 5.2. *Since $\|\Theta\| < 1$, $\boldsymbol{\theta}$ is a contraction mapping and therefore has a unique fixed point.*

6. Existence of eigenfunction

Suppose now that we have found a solution $\boldsymbol{\theta}(x)$ of (4.6), with derivative Θ. The following theorem gives sufficient conditions for obtaining an eigenfunction through the integration of (4.4).

Theorem 6.1. *If $\Gamma_{22}(x, y) + \Gamma_{11}(y, \boldsymbol{\theta}(y)) + \Gamma_{12}(y, \boldsymbol{\theta}(y)) \cdot \Theta(y)$ is a symmetric negative-definite matrix for all $x, y \in \mathbb{R}^n$, then λ and f exist satisfying (2.5).*

Proof. The usual condition that $\nabla_1 A(x, \boldsymbol{\theta}(x))$, quâ function of x, be the gradient of some scalar function f of x requires that $\partial_i((\partial_k A)(x, \boldsymbol{\theta}(x))$ be symmetric in i and k $(i, k = 1,\ldots,n)$ for all x, which is easily found to be equivalent to the symmetry, for all x, of the matrix

$$\Gamma_{11}(x, \boldsymbol{\theta}(x)) + \Gamma_{12}(x, \boldsymbol{\theta}(x)) \cdot \Theta(x).$$

But this is guaranteed by the theorem hypothesis, since Γ_{22} is symmetric.

Hence $f: \mathbb{R}^n \to \mathbb{R}$ exists such that

$$(\nabla_1 A)(x, \boldsymbol{\theta}(x)) = (\nabla f)(x) \quad \forall\, x \in \mathscr{C}, \tag{6.1}$$

so

$$(\nabla_1 A)(\boldsymbol{\theta}(x), \boldsymbol{\theta}^2(x)) = (\nabla f)(\boldsymbol{\theta}(x)) \quad \forall\, x \in \mathscr{C}, \tag{6.2}$$

whence by (4.6)

$$(\nabla_2 A)(x, \boldsymbol{\theta}(x)) + (\nabla f)(\boldsymbol{\theta}(x)) = 0, \tag{6.3}$$

i.e.

$$\nabla_2 (A(x, y) + f(y)) = 0 \quad \text{at } y = \boldsymbol{\theta}(x). \tag{6.4}$$

Moreover, the second derivative of $A(x, y) + f(y)$ with respect to y is easily found, using (6.1), to be exactly the matrix given as negative-definite in the theorem hypothesis. Thus $A(x, y) + f(y)$ is concave in y, so (6.4) shows that

$$A(x, \boldsymbol{\theta}(x)) + f(\boldsymbol{\theta}(x)) = \max_{y \in \mathscr{C}} (A(x, y) + f(y)) \quad \forall\, x \in \mathscr{C}. \tag{6.5}$$

On the other hand, if we take the gradient of

$$A(x, \boldsymbol{\theta}(x)) + f(\boldsymbol{\theta}(x)) - f(x)$$

(use (4.3)) and make use of (6.1), (6.3) we find this gradient is identically zero. Thus

$$A(x, \boldsymbol{\theta}(x)) + f(\boldsymbol{\theta}(x)) - f(x) = \text{const} = \lambda \text{ (say)} \quad \forall\, x \in \mathscr{C}, \tag{6.6}$$

$\mathscr{C}$ is arbitrarily large. So, from (6.5), (6.6), λ and f exist such that (2.5) holds. □

7. The concave-quadratic case

Suppose now that

$$A(x, y) = \tfrac{1}{2}x'Fx + x'Gy + \tfrac{1}{2}y'Hy + c'x + d'y, \tag{7.1}$$

where F, G, H are symmetric matrices, and dashes denote transposition. Relation (4.6) becomes now

$$G\theta^2(x) + (F+H)\theta(x) + Gx + c + d = 0 \quad \forall\, x \in \mathscr{C}, \tag{7.2}$$

and Γ is the constant matrix:

$$\Gamma = \begin{bmatrix} F & G \\ G & H \end{bmatrix}. \tag{7.3}$$

We assume that A is a strictly concave function of its arguments, i.e. that Γ is negative-definite.

Let us temporarily further assume that $(F+H)$ and G are diagonal matrices:

$$\begin{aligned} (F+H) &= \operatorname{diag}(\alpha_i), \quad i = 1,\ldots,n, \\ G &= \operatorname{diag}(\beta_i), \quad i = 1,\ldots,n, \end{aligned} \tag{7.4}$$

where from the negative-definiteness of Γ it follows that F, H and therefore $(F+H)$ are negative-definite, so

$$\alpha_i < 0, \quad i = 1,\ldots,n. \tag{7.5}$$

Also, if $z \in \mathbb{R}^{2n}$ has its ith component equal to 1, its $(n+i)$th component equal to ± 1 and other components zero,

$$z'\Gamma z = \alpha_i \pm 2\beta_i, \quad i = 1,\ldots,n$$

Hence by the negative-definiteness of Γ

$$\alpha_i < -2|\beta_i|. \tag{7.6}$$

Consider the following equation for the scalar y_i:

$$\beta_i y_i^2 + \alpha_i y_i + \beta_i = 0. \tag{7.7}$$

If $\beta_i = 0$, there is a unique root equal to zero. Otherwise there are two roots whose product is unity. Relation (7.6) ensures that they are real and unequal. Hence, in all cases (7.7) has a unique real root p_i (say) with

$$|p_i| < 1. \tag{7.8}$$

Define the matrix P:

$$P = \operatorname{diag}(p_i). \tag{7.9}$$

Then (7.8) implies

$$\|P\| < 1. \tag{7.10}$$

Further define $q \in \mathbb{R}^n$ to have components

$$q_i = \frac{-(c_i + d_i)}{(1+p_i)\beta_i + \alpha_i}, \quad i = 1,\ldots,n, \tag{7.11}$$

where c_i, d_i are the components of $\boldsymbol{c}$, $\boldsymbol{d}$ in (7.2). (The denominator of (7.11) cannot be zero by virtue of (7.6).)

From (7.7), (7.11) we confirm that

$$\beta_i\left(p_i^2 x_i + p_i q_i + q_i\right) + \alpha_i(p_i x_i + q_i) + \beta_i x_i + c_i + d_i \tag{7.12}$$

is identically zero in the arbitrary indeterminate $x_i \in \mathbb{R}$.

But this says that the transformation $\boldsymbol{\theta}$ defined on $\boldsymbol{x} = [x_i] \in \mathbb{R}^n$,

$$\boldsymbol{\theta}(\boldsymbol{x}) = P\boldsymbol{x} + \boldsymbol{q} \tag{7.13}$$

by $x_i \mapsto p_i x_i + q_i$, satisfies (7.2). And (7.10) implies

$$\|\Theta\| < \sigma \quad \text{for some } \sigma < 1. \tag{7.14}$$

Hence, $\boldsymbol{\theta}$ has a fixed point which is evidently

$$\boldsymbol{\xi} = (I - P)^{-1}\boldsymbol{q}. \tag{7.15}$$

Also

$$F + H + GP = \operatorname{diag}(\alpha_i + \beta_i p_i), \quad i = 1,\ldots,n \tag{7.16}$$

and

$$\alpha_i + \beta_i p_i < 0, \quad i = 1,\ldots,n \tag{7.17}$$

by (7.6), (7.8). Hence $F + H + GP$ is (symmetric and) negative-definite, so we may apply Theorem 6.1 to infer the existence of a solution to the eigenfunction-eigenvalue problem.

We may now drop assumption (7.4) that $(F + H)$ and G are diagonal since by the negative-definiteness of $(F + H)$ there exists a nonsingular $(n \times n)$ matrix T which simultaneously diagonalises both $(F + H)$ and G. If the eigenfunction of $A(T\boldsymbol{x}, T\boldsymbol{y})$ derived as above is g, then since $T\boldsymbol{x}$, $T\boldsymbol{y}$ are as general as $\boldsymbol{x}$ and $\boldsymbol{y}$ in (2.1), it is clear that $g \circ T^{-1}$ is an eigenfunction for A. We have evidently established the following.

Theorem 7.1. *If $A(x, y)$ is a strictly concave quadratic function, then the eigenfunction-eigenvalue problem is soluble.*

We remark that the assumption of strict concavity for A may be weakened. Provided that $(F + H)$ and G are simultaneously diagonalisable and $(F + H) + GP$ is negative-definite, it is clear that the given construction will work.

8. Explicit solutions for the concave-quadratic case

With the notation of Section 7 let (7.1) again be a general strictly concave quadratic function. If g is the eigenfunction of $A(T\boldsymbol{x}, T\boldsymbol{y})$, then using (6.1) we have

$$(\nabla g(\boldsymbol{x})) = \boldsymbol{x}'(T'FT + T'GTP) + (\boldsymbol{q}'T'GT + \boldsymbol{c}'T) \tag{8.1}$$

which integrates to

$$g(\boldsymbol{x}) = \tfrac{1}{2}\boldsymbol{x}'(T'FT + T'GTP)\boldsymbol{x} + (T'GT\boldsymbol{q} + T'\boldsymbol{c})'\boldsymbol{x} \tag{8.2}$$

producing for $A(x, y)$ the eigenfunction

$$f(x) = g(T^{-1}x) = \tfrac{1}{2}x'(F + GTPT^{-1})x + (GTq + c)'x. \tag{8.3}$$

If $A(Tx, Ty) + g(y)$ is maximised with respect to y when $y = Px + q$, then $A(x, y) + f(y)$ is maximised with respect to y when

$$y = TPT^{-1}x + Tq. \tag{8.4}$$

Hence,

$$\theta(x) = Qx + r, \tag{8.5}$$

where

$$Q = TPT^{-1} \quad \text{and} \quad r = Tq. \tag{8.6}$$

So (8.3) may be written as

$$f(x) = \tfrac{1}{2}x'(F + GQ)x + (Gr + c)'x. \tag{8.7}$$

The fixed point of θ is

$$\xi = (I - Q)^{-1}r. \tag{8.8}$$

Substituting for $\theta(x)$ from (8.5) in (7.2) we find that Q and r satisfy

$$GQ^2 + (F + H)Q + G = 0, \tag{8.9}$$

$$[G(I + Q) + (F + H)]r = -(c + d). \tag{8.10}$$

Hence, replacing r by $(I - Q)\xi$ in (8.10):

$$(c + d) = -[G - GQ^2 + F + H - (F + H)Q]\xi = -[F + H + 2G]\xi \quad \text{(using (8.9))}. \tag{8.11}$$

Hence, the eigenvalue is

$$\lambda = A(\xi, \xi) = \tfrac{1}{2}\xi'(F + 2G + H)\xi + (c + d)'\xi = \tfrac{1}{2}(c + d)'\xi. \tag{8.12}$$

9. A numerical example

Suppose for A of (7.1) that

$$\Gamma = \begin{bmatrix} -6 & 4 & -2 & 1 \\ 4 & -4 & 1 & -1 \\ -2 & 1 & -3 & 2 \\ 1 & -1 & 2 & -2 \end{bmatrix}; \qquad c = \begin{bmatrix} -1 \\ 0 \end{bmatrix}; \qquad d = \begin{bmatrix} 2 \\ -1 \end{bmatrix}. \tag{9.1}$$

Thus

$$(F + H) = \begin{bmatrix} -9 & 6 \\ 6 & -6 \end{bmatrix}; \qquad G = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}. \tag{9.2}$$

By routine methods we confirm that Γ is negative-definite and that a common diagonalising matrix for $(F+H)$ and G is

$$T=\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}. \tag{9.3}$$

In fact

$$T'(F+H)T=\begin{bmatrix}-3 & 0\\ 0 & -6\end{bmatrix};\qquad T'GT=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}. \tag{9.4}$$

From (7.7), (7.9):

$$P=\begin{bmatrix}\frac{1}{2}(-3+\sqrt{5}) & 0\\ 0 & -3+2\sqrt{2}\end{bmatrix}. \tag{9.5}$$

Recalling that we are working with the *transformed* $\boldsymbol{c}$ and $\boldsymbol{d}$ of $A(T\boldsymbol{x}, T\boldsymbol{y})$, (7.11) gives

$$\boldsymbol{q}=\begin{bmatrix}0\\ \frac{1}{4}(-2+\sqrt{2})\end{bmatrix} \tag{9.6}$$

and the gradient of the eigenfunction for $A(T\boldsymbol{x}, T\boldsymbol{y})$ from (6.1) is

$$\left[\tfrac{1}{2}(-1-\sqrt{5})x_1-1,\ (-1-2\sqrt{2})x_2+\tfrac{1}{4}(2-\sqrt{2})\right], \tag{9.7}$$

which integrates to $g(\boldsymbol{x})$ (say) where

$$g(x_1, x_2)=-\tfrac{1}{4}(1+\sqrt{5})x_1^2-\tfrac{1}{2}(1+2\sqrt{2})x_2^2-x_1+\tfrac{1}{4}(2-\sqrt{2})x_2. \tag{9.8}$$

Then the eigenfunction of A is

$$\begin{aligned} f(\boldsymbol{x}) &= g(T^{-1}\boldsymbol{x})=g(x_1, x_2-x_1)\\ &= -\tfrac{1}{4}(3+4\sqrt{2}+\sqrt{5})x_1^2+(1+2\sqrt{2})x_1x_2\\ &\quad -\tfrac{1}{2}(1+2\sqrt{2})x_2^2-\tfrac{1}{4}(6-\sqrt{2})x_1+\tfrac{1}{4}(2-\sqrt{2})x_2. \end{aligned} \tag{9.9}$$

We find further that

$$Q=\begin{bmatrix}\frac{1}{2}(-3+\sqrt{5}) & 0\\ \frac{1}{2}(3+\sqrt{5}-4\sqrt{2}) & -3+2\sqrt{2}\end{bmatrix}, \tag{9.10}$$

$$\boldsymbol{r}=\boldsymbol{q},$$

$$\boldsymbol{\xi}=\begin{bmatrix}0\\ -\frac{1}{8}\end{bmatrix}, \tag{9.11}$$

$$\lambda=\tfrac{1}{16}. \tag{9.12}$$

10. The separable case

Returning to the general case of a function satisfying the conditions of Section 3, we shall say that A is *separable* if it is of the form

$$A(\boldsymbol{x}, \boldsymbol{y})=R(\boldsymbol{x})+S(\boldsymbol{y}). \tag{10.1}$$

Relation (4.6) becomes now

$$(\nabla R)(\boldsymbol{\theta}(\boldsymbol{x})) + (\nabla S)(\boldsymbol{\theta}(\boldsymbol{x})) = 0. \tag{10.2}$$

So $\boldsymbol{\theta}(\boldsymbol{x})$ is a solution of

$$(\nabla \Xi)(\boldsymbol{t}) = 0, \tag{10.3}$$

where

$$\Xi(\boldsymbol{t}) = R(\boldsymbol{t}) + S(\boldsymbol{t}). \tag{10.4}$$

If Ξ is strictly concave, then (10.3) has one solution at most, so $\boldsymbol{\theta}(x)$ is constant and equal to this solution, which then is clearly the fixed point $\boldsymbol{\xi}$. Relation (6.1) is

$$(\nabla f)(\boldsymbol{x}) = (\nabla R)(\boldsymbol{x}), \tag{10.5}$$

whence, taking a zero constant of integration,

$$f = R, \tag{10.6}$$

then

$$A(\boldsymbol{x}, \boldsymbol{y}) + f(\boldsymbol{y}) = R(\boldsymbol{x}) + \Xi(\boldsymbol{y}), \tag{10.7}$$

which for all $\boldsymbol{x}$ is maximised with respect to $\boldsymbol{y}$ by $\boldsymbol{y} = \boldsymbol{\xi}$ because Ξ is concave. Hence, using (10.6), (10.7):

$$\max_{\boldsymbol{y}} (A(\boldsymbol{x}, \boldsymbol{y}) + f(\boldsymbol{y})) = R(\boldsymbol{x}) + \Xi(\boldsymbol{\xi}) = f(\boldsymbol{x}) + \Xi(\boldsymbol{\xi}).$$

So in this case $\lambda = \Xi(\boldsymbol{\xi})$, i.e.

$$\lambda = R(\boldsymbol{\xi}) + S(\boldsymbol{\xi}) = A(\boldsymbol{\xi}, \boldsymbol{\xi}). \tag{10.8}$$

11. The orbit problem

Let A again be any function satisfying the assumptions listed in Section 3.

Consider now the problem mentioned in Section 2, of constructing a sequence of transitions, starting at an arbitrary point $\boldsymbol{x} \in \mathscr{C}$, so as to arrive at the self-corresponding point $\boldsymbol{\xi}$. A transition from $\boldsymbol{y} \in \mathscr{C}$ to $\boldsymbol{z} \in \mathscr{C}$ gives profit $A(\boldsymbol{y}, \boldsymbol{z})$ and the profit of the sequence is the sum (if well-defined) of the profits of the transitions. Find a strategy to reach $\boldsymbol{\xi}$ from each $\boldsymbol{x} \in \mathscr{C}$ at maximal profit.

One representation of such a problem would be as follows. Let Δ be the class of all functions $\psi : \mathbf{R}^n \to \mathbf{R}^n$ such that

$$\psi(\mathscr{C}) \subseteq \mathscr{C}, \tag{11.1}$$

and for all $\boldsymbol{x} \in \mathscr{C}$ the sequence $\boldsymbol{x}, \psi(\boldsymbol{x}), \psi(\psi(\boldsymbol{x})), \ldots$ converges to $\boldsymbol{\xi}$. (11.2)

We shall call the elements of Δ *strategies*, and the sequence of transitions a *path* from $\boldsymbol{x}$ to $\boldsymbol{\xi}$.

If conditions (5.1) hold, then by Theorem 5.1 and the fixed-point theorem

$$\boldsymbol{\theta} \in \Delta. \tag{11.3}$$

Define $\psi^r(\boldsymbol{x})$ by

$$\psi^r(\boldsymbol{x}) = \psi(\psi^{(r-1)}(\boldsymbol{x})), \quad r = 2, 3, \ldots. \tag{11.4}$$

Note that (2.5) implies

$$A(x,y)+f(y)\leqslant\lambda+f(x)\quad\forall\, x,y\in\mathscr{C}. \tag{11.5}$$

Thus

$$\begin{aligned}&A(x,\psi(x))+f(\psi(x))\leqslant\lambda+f(x),\\&A(\psi,(x),\psi^2(x))+f(\psi^2(x))\leqslant\lambda+f(\psi(x)),\\&A(\psi^{r-1}(x),\psi^r(x))+f(\psi^r(x))\leqslant\lambda+f(\psi^{r-1}(x)),\end{aligned} \tag{11.6}$$

whence

$$\sum_{k=1}^{r}A(\psi^{k-1}(x),\psi^k(x))\leqslant r\lambda+f(x)-f(\psi^r(x)), \tag{11.7}$$

where $\psi^0(x)$ means x and $\psi^1(x)$ means $\psi(x)$. However, when ψ is $\boldsymbol{\theta}$, inequalities (11.6) and thus (11.7) hold with equality. Hence for all $\psi\in\Delta$

$$\begin{aligned}\limsup_{r\to\infty}\sum_{k=1}^{r}A(\psi^{k-1}(x),\psi^k(x))&\leqslant\sum_{k=1}^{\infty}A(\boldsymbol{\theta}^{k-1}(x),\boldsymbol{\theta}^k(x))\\&=\begin{cases}-\infty, & \text{if } \lambda<0,\\ f(x)-f(\boldsymbol{\xi}), & \text{if } \lambda=0,\\ +\infty, & \text{if } \lambda>0.\end{cases}\end{aligned} \tag{11.8}$$

We have thus proved the following result.

Theorem 11.1. *If, and only if, $\lambda=0$, there exists a strategy such that the total profit of the path from any x to $\boldsymbol{\xi}$ is finite and well-defined and not less than the total profit for any other strategy for which the total profit is well-defined.*

We may remark that if we choose the constant of integration such that $f(\boldsymbol{\xi})=0$, then (11.8) shows that the eigenfunction $f(x)$ gives the optimal profit of a path from each x to $\boldsymbol{\xi}$.

References

[1] R.A. Cuninghame-Green, *Minimax Algebra*, Lecture Notes in Economics and Mathematical Systems **166** (Springer-Verlag, Berlin, 1979).

[2] U. Zimmermann, *Linear and Combinatorial Optimization in Ordered Algebraic Structures*, Annals of Discrete Mathematics **10** (North-Holland, Amsterdam, 1981).

[3] A. Brøndsted, Convexification of conjugate functions, *Math. Scand.* **36** (1975) 131–136.

[4] J.J. Moreau, Inf-convolution, sous-additivité, convexité des fonctions numériques, *J. Math. Pures Appl.* **49** (1970) 109–154.

[5] R. Bellman and W. Karush, On a new functional transform in analysis: the maximum transform, *Bull. Amer. Math. Soc.* **67** (1961) 501–503.

[6] L. McLinden, An extension of Fenchel's duality theorem to saddle functions and dual minimax problems, *Pacific J. Math.* **50** (1974) 135–158.